

# STABLE W-LENGTH

DANNY CALEGARI AND DONGPING ZHUANG

**ABSTRACT.** We study stable  $W$ -length in groups, especially for  $W$  equal to the  $n$ -fold commutator  $\gamma_n := [x_1, [x_2, \dots [x_{n-1}, x_n] \dots]]$ . We prove that in any perfect group, for any  $n \geq 2$  and any element  $g$ , the stable commutator length of  $g$  is at least as big as  $2^{2-n}$  times the stable  $\gamma_n$ -length of  $g$ . We also establish analogues of Bavard duality for words  $\gamma_n$  and for  $\beta_2 := [[x, y], [z, w]]$ . Our proofs make use of geometric properties of the asymptotic cones of verbal subgroups with respect to bi-invariant metrics. In particular, we show that for suitable  $W$ , these asymptotic cones contain certain subgroups that are *normed vector spaces*.

## 1. INTRODUCTION

Geometric group theory aims to produce functors from the algebraic category of groups and homomorphisms to geometric categories of spaces and structure-preserving maps. The category of metric spaces and isometries does not have enough morphisms for many applications, so one instead typically studies functors from the category of groups and homomorphisms, to the category of metric spaces and 1-Lipschitz (i.e. distance decreasing) maps.

A rich source of such functors arise in the theory of *bounded cohomology*, introduced systematically by Gromov in [12]. In that theory, the metric spaces are usually *normed vector spaces*, and the morphisms *bounded linear operators*. Another rich source of such examples comes from the study of (conjugation-invariant) *norms*; see e.g. [3] for a discussion, and [14] for an application to the theory of mapping class groups.

Such functors are often useful for the study of groups as dynamical objects, where the functor “geometrizes” the group action, and allows one to obtain *a priori* control of dynamical quantities from algebra. For example, *stable commutator length* (hereafter scl; see [6] or [1] for an introduction) has rich connections to 2-dimensional dynamics, symplectic geometry, hyperbolic geometry, and so on.

A natural class of characteristic norms (those invariant under any automorphism), with good monotonicity properties, arise from the theory of *words*. Given a subset  $W$  of a free group  $F$ , a  $W$ -word in  $G$  is the image of some  $w \in W$  under some homomorphism  $\phi : F \rightarrow G$ . The  $W$ -words in  $G$  generate a so-called *verbal subgroup*  $G_W$  (see e.g. [17]), and the  $W$ -length of any  $g \in G_W$  is defined to be the smallest number of  $W$ -words and their inverses in  $G$  whose product is equal to  $g$ . For example, many authors study *square length*, which is  $W$ -length in the case  $W$  consists of the single word  $x^2$ .  $W$ -length (usually for  $W$  consisting of a single word  $w$ ) has been intensively studied in finite groups, and recently some very strong theorems have been obtained by Shalev and his collaborators (e.g. [20, 21, 15]). See

[19] for a survey, and an introduction to some of these methods. However, with some exceptions (notably [18]),  $W$ -length has not been widely studied in (general) infinite groups except in the special case of  $W = \{[x, y]\}$  — i.e. commutator length. Part of the problem is that  $W$ -length seems to be such an unstructured quantity in general, and is exceedingly hard to compute, or even to estimate — even in finite groups! Therefore in this paper we propose to study a suitable “rationalization” of this quantity, namely *stable  $W$ -length*, where the stable  $W$ -length of an element  $g \in G_W$  is defined to be the limit of the quantity ( $W$ -length of  $g^n$  divided by  $n$ ) as  $n \rightarrow \infty$ . Our aim is to generalize (to the extent that it is possible) some of what is known about stable commutator length to more general classes of norms.

Perhaps the most significant difficulty in generalizing the theory of (stable) commutator length to more general stable lengths is that of *linearizing* the theory. The commutator calculus exhibits an intimate connection between the algebraic theory of commutators and the linear theory of (2-dimensional) homology. Commutators arise as boundaries in group homology, and one may obtain a duality (known as *Bavard Duality*; see [1] and [6], Chapter 2) between stable commutator length and certain natural group 1-cocycles called *homogeneous quasimorphisms*. There is no natural homology theory available for stable  $W$ -length for more general  $W$ , but one may obtain a generalization of the theory for commutators by a *geometric* construction. Given a group  $G$ , one considers the Cayley graph of  $[G, G]$  taking as generators *all commutators in  $G$* . This is a metric space (a graph), and the group  $[G, G]$  acts on itself by metric isometries. Because every commutator has “length 1” in this graph, the action of the group  $[G, G]$  on itself is “almost” commutative. When one replaces the Cayley graph by its asymptotic cone (an infinite re-scaled version), the limit becomes *exactly* commutative, and one obtains a normed vector space, which can be identified with one of the natural spaces obtained from the homological approach. It is this geometric construction that generalizes: given any  $W$ , one considers the Cayley graph of  $G_W$  with all  $W$ -words and their inverses as generators. Since  $W$ -words act with bounded length in the Cayley graph, the rescaled asymptotic cone obeys the “law” that all  $W$ -words are *trivial*. With more work, one can obtain a certain subset of the asymptotic cone where, for suitable  $W$ , the resulting group is actually *abelian*, and is in fact a normed vector space where one can establish the analogue of Bavard duality. This program is most successful when  $W$  is an  $n$ -fold commutator  $\gamma_n$ , in which case we are able to establish (for perfect groups  $G$ ) a precise analog of Bavard duality, and to prove the existence of *two-sided* estimates of stable  $\gamma_n$ -length in terms of stable commutator length. This comparison theorem is a genuinely stable phenomenon, and does not hold for “naive”  $\gamma_n$ -length (for a precise statement of results, see § 1.1).

**1.1. Statement of results.** We now give a precise statement of the main results in the paper.

In § 2 we introduce terminology, and establish basic properties. We show that there are various inequalities relating  $W$ -length and stable  $W$ -length in various groups and for various different  $W$ . We use the notation (here and elsewhere) of  $l(*|W)$  for  $W$ -length, and  $sl(*|W)$  for stable  $W$ -length, with  $cl(*)$  and  $scl(*)$  for the special case of commutator length and stable commutator length respectively.

We give simple examples of groups where stable  $w$ -length is nontrivial for essentially all words (free groups, hyperbolic groups) and where stable  $w$ -length is trivial

for all  $w$  ( $\text{SL}(2, A)$  for certain rings of algebraic integers  $A$ ). We also derive some nontrivial estimates (usually upper bounds) on  $\text{sl}$  in free groups. For example,

**(Cor. 2.17)  $\text{sl}(w|w)$  inequality.** *For any  $w$  there is an inequality*

$$1/2 \leq \frac{\text{scl}(w)}{\text{scl}(w) + 1/2} \leq \text{sl}(w|w) \leq 1$$

We strengthen the upper bound by an explicit construction, for certain natural classes of words. Let  $\gamma_n$  denote the  $n$ -fold commutator, so  $\gamma_1 = x$ ,  $\gamma_2 = [x, y]$ ,  $\gamma_3 = [x, [y, z]]$  and so on. We obtain the estimate

**(Prop. 2.28)  $\gamma_n$  inequality.** *For any  $n$  there is an inequality*

$$\text{sl}(\gamma_n|\gamma_n) \leq 1 - 2^{1-n}$$

§ 3 is the heart of the paper. To any group  $G$  and any  $W$ , we study the asymptotic geometry of the Cayley graph  $C_W$  of  $G_W$  with all  $W$ -words as generators. A limit of rescalings of  $C_W$  gives rise to the *asymptotic cone*  $\hat{A}_W$ . The cone  $\hat{A}_W$  is itself a group, with a bi-invariant metric, obeying a nontrivial law. It contains a canonical contractible subgroup  $A_W$  which metrically encodes the values of stable  $W$ -length on  $G_W$ . Our results are most definitive for certain specific  $W$ ; in particular,

**(Thm. 3.11).** *Suppose  $A_W$  is nilpotent (for example, if  $W = \gamma_n$  for some  $n$ ). Then it is a normed vector space (in particular it is abelian).*

Moreover, if  $\beta_2$  denotes the word  $[[x, y], [z, w]]$ ,

**(Thm. 3.14).** *Suppose in some group  $G$  that stable commutator length vanishes identically. Then  $A_W$  is abelian (and hence a normed vector space) for  $W = \beta_2$ .*

The fact that these asymptotic cones are normed vector spaces lets one study their geometry (and thereby stable  $W$ -length) dually, via 1-Lipschitz homomorphisms  $A_W \rightarrow \mathbb{R}$ . Motivated by Bavard duality, we say a function  $\phi : G \rightarrow \mathbb{R}$  is a *weak  $\gamma_n$ -hoq* (“homogeneous quasimorphism”) if it is homogeneous (i.e.  $\phi(g^n) = n\phi(g)$ ) and if there is a least non-negative real number  $D(\phi)$  (called the *defect*) such that for any  $g, h \in G$  there is an inequality

$$|\phi(g) + \phi(h) - \phi(gh)| \leq D(\phi) \min(l(g|\gamma_{n-1}), l(h|\gamma_{n-1}), l(gh|\gamma_{n-1}))$$

**(Thm. 3.23)  $\gamma_n$ -Duality theorem.** *For any perfect group  $G$ , and any  $g \in G$  there is an inequality*

$$\sup_{\phi} \phi(g)/D(\phi) \geq \text{sl}(g|\gamma_n) \geq \sup_{\phi} \phi(g)/2D(\phi)$$

where the supremum is taken over all weak  $\gamma_n$ -hoqs.

From this we derive the following rather surprising estimate, comparing stable commutator length with stable  $\gamma_n$  length in any perfect group:

**(Thm. 3.25)  $\gamma_n$ -Comparison theorem.** *For any perfect group  $G$ , for any  $n \geq 2$  and for any  $g \in G$  there is an inequality*

$$2^{n-2} \text{scl}(g) \geq \text{sl}(g|\gamma_n) \geq \text{scl}(g)$$

Together the  $\gamma_n$ -Comparison theorem and the  $\gamma_n$ -Duality theorem show that stable commutator length in a perfect group  $G$  can be bounded from below by weak  $\gamma_n$ -hoqs. It would be interesting to try to find naturally occurring examples

of such functions arising from realizations of  $G$  as a group of automorphisms of some geometric object.

In a similar vein, we say  $\phi : G \rightarrow \mathbb{R}$  is a *weak  $\beta_2$ -hoq* if it is homogeneous, and if there is a least non-negative real number  $D(\phi)$  (called the *defect*) such that for any  $g, h \in G$  there is an inequality

$$|\phi(g) + \phi(h) - \phi(gh)| \leq D(\phi) \min(l(g|\beta_2), l(h|\beta_2), l(gh|\beta_2))$$

**(Thm. 3.31)  $\beta_2$ -Duality theorem.** *For any perfect group  $G$  in which  $\text{scl}$  vanishes identically, and for any  $g \in G$  there is an inequality*

$$\sup_{\phi} 2\phi(g)/D(\phi) \geq \text{sl}(g|\beta_2) \geq \sup_{\phi} \phi(g)/2D(\phi)$$

where the supremum is taken over all weak  $\beta_2$ -hoqs.

Note that stable commutator length is known to vanish identically in many important classes of groups, including

- amenable groups
- lattices in higher rank Lie groups [4]
- groups of PL homeomorphisms of the interval [5]
- groups that satisfy a law [7]

It is not known whether stable  $W$ -length for various  $W$  vanish in these classes of groups. One of the aims of this paper is to develop tools to approach this question.

Finally, in § 4, we give an elementary construction, using hyperbolic geometry, of a class of elements  $w_n$  in free groups of various ranks for which  $\text{scl}(w_n) = 1/2$  and  $\text{sl}(w_n|\gamma_3) = 1$ , but for which  $\text{cl}(w_n) = 1$  and  $l(w_n|\gamma_3) \geq 2n/3$ .

**1.2. Acknowledgments.** We would like to thank Frank Calegari, Benson Farb, Denis Osin and Dan Segal for helpful comments. We would also like to thank the anonymous referee for a very careful reading and for catching several errors. Danny Calegari was supported by NSF grant DMS 1005246.

## 2. BASIC PROPERTIES

The purpose of this section is to standardize definitions and notation, and to survey some elementary constructions in the theory.

**Notation 2.1.** We use the notation  $[x, y]$  for  $xyx^{-1}y^{-1}$  and  $x^y$  for  $yx y^{-1}$ . With this convention, exponentiation obeys  $(x^y)^z = x^{zy}$ . We also use the notation  $x^*$  for some (unspecified) conjugate of  $x$ .

The following identities, though elementary, are useful.

**Lemma 2.2.** *The following identities hold (the letters denote free generators):*

- (1)  $[x, y] = [y^x, x^{-1}]$
- (2)  $[x, y]^{-1} = [y, x] = [x^y, y^{-1}]$
- (3)  $[x, yz] = [x, y][x, z]^y = [x, y][x, z][[z, x], y]$
- (4)  $[xy, z] = [x, z][y, z]^{x^z} = [x, z][y, z][[z, y], x^z]$
- (5)  $[[y, x], z^x][[x, z], y^z][[z, y], x^y] = 1$

Bullet (5) is known as the *Hall-Witt identity*.

### 2.1. Verbal subgroups.

**Definition 2.3** (*W-word*). Let  $F$  be a free group, and  $W$  a subset of  $F$ . A *W-word* in a group  $G$  is the image of some  $w \in W$  under a homomorphism  $F \rightarrow G$ .

**Definition 2.4** (*verbal subgroup*). Let  $G$  be a group. A *verbal subgroup* of  $G$  is the subgroup generated by the  $W$ -words, for some  $W \subset F$ . The  $W$ -verbal subgroup of  $G$  is denoted  $G_W$ .

The group  $G$  is said to *obey the law W* if  $G_W$  is trivial. A group is said to *obey a law* if  $G_W$  is trivial for some nonempty  $W$ .

For any sets  $V, W \subset F$  with  $F_V = F_W$  (for example, if  $W \subset V \subset F_W$ ) and for any group  $G$ , the verbal subgroups  $G_V$  and  $G_W$  are equal, though the  $V$ -words and the  $W$ -words in  $G$  might not be equal *as sets*.

Since the image of a  $W$ -word under any automorphism is also a  $W$ -word, verbal subgroups are characteristic (i.e. invariant under every automorphism). In particular, a conjugate of a  $W$ -word is a  $W$ -word. In a free group, it is tautologically true that characteristic subgroups are verbal.

In the sequel we are interested in verbal subgroups  $G_W$  where  $W$  is a single element of  $F$ . For the sake of legibility, we use a lower case  $w$  to denote a single element of  $F$ , and likewise denote the  $w$ -subgroup of  $G$  by  $G_w$ . In fact, we are interested in certain explicit words  $w$ . The main  $w$  of interest in this paper are the *n-fold commutators*.

**Definition 2.5.** Here and in the sequel, we let  $\gamma_n$  denote the  $n$ -fold commutator of the generators of  $F_n$ . That is,

$$\gamma_1 = x, \quad \gamma_2 = [x, y], \quad \gamma_3 = [x, [y, z]], \quad \gamma_4 = [x, [y, [z, w]]]$$

and so on.

**Definition 2.6.** A word  $w \in F$  is *reflexive* if  $w^{-1}$  is a  $w$ -word in  $F$ .

**Lemma 2.7.** *The words  $\gamma_n$  are all reflexive. Moreover, any nested bracket of  $n$  elements is a  $\gamma_n$  word.*

*Proof.* The second statement means, for example, that an expression like  $[[x, y], z]$  is a  $\gamma_3$  word, that  $[x, [[[w, [u, v]], z], y]]$  is a  $\gamma_5$  word, and so on. This and the reflexivity of  $\gamma_n$  follow inductively from Lemma 2.2, bullet (2).  $\square$

For any  $n$  and any  $G$ , the subgroups  $G_{\gamma_n}$  are usually denoted  $G_n$  and called the *lower central series*. The successive quotients  $G/G_n$  are (universal)  $(n-1)$ -step nilpotent quotients of  $G$  (so that for instance  $G/G_2$  is abelian,  $G/G_3$  is 2-step nilpotent, and so on).

**Definition 2.8.** Here and in the sequel, we define  $\beta_1 = [x, y]$ , and for  $n > 1$  we let  $\beta_n$  denote the commutator of two copies of  $\beta_{n-1}$  (in different generating sets). Hence

$$\beta_1 = [x, y], \quad \beta_2 = [[x, y], [z, w]], \quad \beta_3 = [[[x, y], [z, w]], [[s, t], [u, v]]]$$

and so on.

**Lemma 2.9.** *The words  $\beta_n$  are reflexive.*

*Proof.* This follows inductively from Lemma 2.2, bullet (2).  $\square$

The subgroups  $G_{\beta_n}$  are usually denoted  $G^{(n)}$  and called the *derived series*. The successive quotients  $G/G^{(n)}$  are (universal)  $n$ -step solvable quotients of  $G$ . Note that  $\beta_1 = \gamma_2$ .

## 2.2. $W$ -length and $w$ -length.

**Definition 2.10** ( $W$ -length). Let  $W \subset F$  be given. Let  $G$  be a group, and let  $G_W$  be the subgroup generated by  $W$ -words. If  $g \in G_W$ , the  $W$ -length of  $g$ , denoted  $l(g|W)$ , is the minimum number of  $W$ -words and their inverses in  $G$  whose product is  $g$ , and the *stable  $W$ -length* of  $g$ , denoted  $sl(g|W)$ , is the limit

$$sl(g|W) = \lim_{n \rightarrow \infty} \frac{l(g^n|W)}{n}$$

If  $W$  consists of a single element, we usually denote it by a lower case  $w$ , and define  $w$ -length and stable  $w$ -length. We are typically concerned with this case in the sequel.

If  $w$  is reflexive (which shall usually be the case in the sequel), the inverse of a  $w$ -word is a  $w$ -word. By abuse of notation we will sometimes refer to both  $w$ -words and their inverses as  $w$ -words. The meaning should be clear from context in each case.

**Notation 2.11.** If we need to emphasize that  $g$  is in specific group  $G$ , we use the notation  $l_G(g|W)$  and  $sl_G(g|W)$ .

In the special case that  $w = [x, y]$  in the free group generated by  $x$  and  $y$ , we refer to  $w$ -length as *commutator length*, and stable  $w$ -length as *stable commutator length*, and denote them by  $cl(\cdot)$  and  $sl(\cdot)$  respectively.

**Notation 2.12.** We abbreviate the product of  $n$  (arbitrary)  $w$ -words by  $w^{\{n\}}$  (or  $W^{\{n\}}$  for  $W$ -words). This should not be confused with  $w^n$ , the  $n$ th power of a specific word  $w$ .

**Lemma 2.13.** *Suppose  $w$  is not in  $[F, F]$ . Then  $sl(\cdot|w)$  is identically zero in every group.*

*Proof.* Since  $w$  is not in  $[F, F]$ , there is a surjective homomorphism  $F \rightarrow \mathbb{Z}$  sending  $w$  to some nonzero  $n$ . Without loss of generality, we may assume  $n > 0$ . But then in any group  $l(g^{nm}|w) \leq 1$  so  $sl(g|w) = 0$ .  $\square$

**Lemma 2.14.** *The functions  $l(\cdot|W)$  and  $sl(\cdot|W)$  are characteristic, and their values do not increase under homomorphisms.*

*Proof.* The group  $\text{Aut}(G)$  permutes the canonical generators of  $G_W$ , proving the first claim. The image of a  $W$ -word under any homomorphism is a  $W$ -word, proving the second claim.  $\square$

By convention, we set  $sl(g|W) = sl(g^n|W)/n$  whenever  $g^n \in G_W$  (even if  $g$  is not necessarily in  $G_W$ ).

**2.3. Inequalities.** If  $w$  and  $v$  are contained in free groups  $F$  and  $F'$  respectively, by abuse of notation we think of  $w$  and  $v$  as being contained in the single free group  $F * F'$ . By convention therefore we think of all abstract words as being contained in a single (infinitely generated) free group, which we denote hereafter by  $F$ .

**Lemma 2.15** (fundamental inequality). *For any  $v, w \in F$  and any  $g \in G$  there is an inequality*

$$l_G(g|v) \leq l_G(g|w)l_F(w|v)$$

*Proof.* Write  $g$  minimally as a product of  $w$ -word or their inverses, and then rewrite each of these as a minimal product of  $v$ -words or their inverses. This gives an expression for  $g$  as a product of  $v$ -words and their inverses.  $\square$

Note that for this lemma to make sense we must have  $g \in G_w$  and  $w \in F_v$ . In the case  $v = [x, y]$  we obtain  $cl_G(g) \leq l_G(g|w)cl_F(w)$ . In the case of stable commutator length, one obtains a better estimate

**Lemma 2.16** (scl inequality). *For any  $w \in F$  and any  $g \in G$  and  $n \in \mathbb{Z}$  there is an inequality*

$$scl_G(g^n) \leq l_G(g^n|w)scl_F(w) + \frac{1}{2}(l_G(g^n|w) - 1)$$

Moreover,

$$scl_G(g) \leq sl_G(g|w)(scl_F(w) + \frac{1}{2})$$

*Proof.* Express  $g^n$  as a product  $w_1 \cdots w_m$  of  $w$ -words or their inverses, where  $m = l_G(g^n|w)$ . For any even  $k$  we have

$$g^{nk} = (w_1 \cdots w_m)^k = w_1^k \cdots w_m^k v$$

where  $v$  is a product of at most  $(m-1)k/2$  commutators (see e.g. [6], Lemma 2.24 for the case  $m = 2$ , and apply induction). Taking the limit as  $k \rightarrow \infty$  we obtain the first inequality.

Taking the limit  $n \rightarrow \infty$  and stabilizing gives us the second inequality.  $\square$

**Corollary 2.17** ( $sl(w|w)$  inequality). *For any  $w \in F$  there is an inequality*

$$1/2 \leq \frac{scl(w)}{scl(w) + 1/2} \leq sl(w|w) \leq 1$$

*Proof.* The first inequality follows from the fact that  $scl(w) \geq \frac{1}{2}$  for any nontrivial  $w$  in a free group (see e.g. [6], Theorem 4.111). The second inequality follows from Lemma 2.16 by setting  $g = w$ . The last inequality follows from the tautological expression of  $w^n$  as a product of  $n$  copies of  $w$ .  $\square$

A natural question to ask is the following:

**Question 2.18** (strict inequality). *For which  $w \in F$  is there is a strict inequality*

$$sl(w|w) < 1 \text{ ?}$$

This question has a positive answer for several interesting classes of words.

*Example 2.19.* For each  $g$ , let  $x_1, y_1, \dots, x_g, y_g$  be free generators, and let  $w_g = [x_1, y_1] \cdots [x_g, y_g]$ . Since  $scl(w_g) = g - \frac{1}{2}$  (see e.g. Bavard [1]) we obtain a lower bound  $sl(w_g|w_g) \geq 1 - 1/2g$ . On the other hand, an expression of  $w_g^n$  as a product of  $n(g-1/2) + o(n)$  commutators is also an expression as a product of  $n(g-\frac{1}{2})/g + o(n)$   $w_g$ -words. So

$$sl([x_1, y_1] \cdots [x_g, y_g] | [x_1, y_1] \cdots [x_g, y_g]) = 1 - 1/2g$$

From the definition,  $sl(g|w) \leq l(g^n|w)/n$  for any  $n$ . On the other hand, this inequality can often be definitely improved:

**Lemma 2.20.** *For any  $w \in F$  and any  $g \in G$ , there is an inequality*

$$\mathrm{sl}_G(g|w) \leq (l(g^n|w) - 1 + \mathrm{sl}(w|w))/n$$

*Proof.* Suppose  $g^n = w_1 \cdots w_m$  where  $m = l(g^n|w)$ . For any  $k$  we can write

$$\begin{aligned} g^{nk} &= w_1 \cdots w_m w_1 \cdots w_m \cdots w_1 \cdots w_m \quad (mk \text{ terms}) \\ &= w_1^k w_1' w_2' \cdots w_{(m-1)k}' \end{aligned}$$

where each  $w_i'$  is a  $w$ -word (since  $xy = y(y^{-1}xy)$ , and any conjugate of a  $w$ -word or its inverse is a  $w$ -word or its inverse). In particular, if  $k$  is large, then  $w_1^k$  can be expressed as a product of  $k \mathrm{sl}(w|w) + o(k)$   $w$ -words and therefore  $g^{nk}$  can be expressed as a product of  $k \mathrm{sl}(w|w) + (m-1)k + o(k)$   $w$ -words. Taking  $k$  large, we obtain the desired inequality.  $\square$

**Corollary 2.21.** *If  $l(w^n|w) = m$  then  $\mathrm{sl}(w|w) \leq (m-1)/(n-1)$ .*

*Proof.* Substitute  $w = g$  in Lemma 2.20.  $\square$

We end this section with a couple of examples, illustrating the range of possibilities one can expect in certain specific classes of groups.

*Example 2.22* (Free group). If  $F$  is a free group,  $\mathrm{scl}(g)$  is positive for every  $g \in [F, F]$  (in fact,  $\mathrm{scl}(g) \geq 1/2$ ). So for every  $w$  with  $F_w \subset [F, F]$  we get  $\mathrm{sl}_F(g|w) \geq \mathrm{scl}_F(g)/(\mathrm{scl}_F(w) + 1/2) > 0$ . In other words, for every  $w$  for which  $\mathrm{sl}(*|w)$  is not identically zero in every group, it is positive on every element in a free group. A similar phenomenon holds in groups for which  $\mathrm{scl}$  is typically positive, such as hyperbolic groups (see e.g. [6], Chapter 3).

*Example 2.23* ( $\mathrm{SL}(2, A)$ ). Carter-Keller-Paige [23] Thm. 6.1. show that if  $A$  is the ring of integers in a number field  $K$  containing infinitely many units, and if  $T$  is an element of  $\mathrm{SL}(2, A)$  which is not a scalar matrix (i.e. not of the form  $\lambda \cdot \mathrm{id}$ ) then  $\mathrm{SL}(2, A)$  has a finite index normal subgroup which is boundedly generated by conjugates of  $T$ . That is, in this finite index normal subgroup, every element can be written as a bounded number of conjugates of  $T$ .

For any  $w$ , one can find  $T$  in  $\mathrm{SL}(2, A)$  which is a  $w$ -word but not a scalar matrix (to see this, observe that  $\mathrm{SL}(2, A)$  contains  $\mathrm{SL}(2, \mathbb{Z})$  which contains a free group, none of whose nontrivial elements is a scalar matrix). It follows that  $\mathrm{SL}(2, A)$  has a finite index subgroup in which every element has finite  $w$ -length. Consequently  $\mathrm{sl}(*|w)$  vanishes identically in  $\mathrm{SL}(2, A)$  for every  $w$ .

A similar phenomenon holds in groups such as  $\mathrm{SL}(n, \mathbb{Z})$  when  $n \geq 3$ .

**2.4. Culler's identity.** Culler [10], in his analysis of commutator length in free groups, discovered the beautiful identity

$$[x, y]^3 = [y^x, x^{y^{-1}} x^{-2}] [x^{y^{-1}}, y^2]$$

This identity expresses algebraically the geometric fact that a once-punctured torus admits an (irregular) cover of degree 3 with one boundary component. We can get a lot of mileage out of this identity; this is not really for a very specific reason, rather because this is the simplest identity which certifies that  $\mathrm{scl}([x, y]) < 1$ , and we are able to “bootstrap” this identity to show  $\mathrm{sl}(w|w) < 1$  for many  $w$ .



*Example 2.24*  $([x, y^n])$ . If we set  $w = [x, y^n]$  then we obtain  $l(w^3|w) \leq 2$  and therefore  $sl(w|w) \leq \frac{1}{2}$ . Together with Corollary 2.17 this implies

$$sl([x, y^n]|[x, y^n]) = \frac{1}{2}$$

*Example 2.25*  $([x, y]^2)$ . If we denote  $a = [y^x, x^{y^{-1}}x^{-2}]$  and  $b = [x^{y^{-1}}, y^2]$  then Culler's identity says

$$\begin{aligned} ([x, y]^2)^6 &= ([x, y]^3)^4 = abababab \\ &= aabbcaabbc \text{ for } c = [b^{-1}, b^{-1}a^{-1}b^{-1}] \\ &= a^2b^2c^2(a')^2(b')^2 \end{aligned}$$

for suitable commutators  $a', b'$ . Hence  $l([x, y]^2|[x, y]^2) \leq 5$  and therefore

$$\frac{2}{3} \leq sl([x, y]^2|[x, y]^2) \leq \frac{4}{5}$$

**2.5.  $\gamma_n$  inequalities.** Generalizations of Culler's identity can be obtained from a uniform topological argument.

**Lemma 2.26** (generalized Culler identity). *For all  $x$ , for any  $n$  and any  $y$ , there is an identity of the form*

$$[x, y]^{2n+1} = [* , y^{n+1}]^* \cdot \text{product of } n \text{ words of the form } [* , y]^*$$

*Proof.* Let  $F = \langle x, y \rangle$ . For each  $n$ , we construct a permutation representation  $\rho_n : F \rightarrow S_{2n+1}$ . We think of an element of  $S_{2n+1}$  as a bijective function from  $\{0, 1, \dots, 2n\}$  to itself, and denote such a function by an ordered string of values. With this convention, we define

$$\begin{aligned} \rho_n(x) &= 2n \ 2n-1 \ 2n-2 \ \dots \ 1 \ 0 \\ \rho_n(y) &= 1 \ 2 \ 3 \ \dots \ n \ 0 \ n+1 \ \dots \ 2n \end{aligned}$$

In words,  $\rho_n(y)$  cycles the first  $n+1$  elements and fixes the rest, while  $\rho_n(x)$  is the involution which reverses the order.

If we think of  $F$  as the fundamental group of a once-punctured torus  $S$ , then  $x$  and  $y$  can be represented by embedded essential loops  $\alpha$  and  $\beta$  intersecting in one point. Since  $\beta$  is homologically essential, it is non-separating. Let  $S'$  be the  $2n+1$ -fold cover of  $S$  whose monodromy representation is  $\rho_n$ . By the definition of  $\rho_n$ , the permutation  $\rho_n([x, y])$  is a single  $2n+1$ -cycle, and therefore  $\partial S'$  consists of a single boundary component, mapping to  $\partial S$  with degree  $2n+1$  under the covering projection.

Likewise,  $\beta$  has  $n+1$  preimages; one (call it  $\beta_0$ ) maps to  $\beta$  with degree  $n+1$ . The others  $\beta_i$  for  $1 \leq i \leq n$  map with degree 1. We claim that there are disjoint, embedded curves  $\alpha_i$  in  $S'$  so that each  $\alpha_i$  and  $\beta_i$  intersect transversely in one point, and  $\alpha_i$  does not intersect  $\beta_j$  for  $i \neq j$ . Evidently the lemma follows from this.

To prove this claim we first show that  $\cup \beta_i$  is non-separating. Let  $T$  be a component of  $S' - \cup \beta_i$ . Then  $T$  covers  $S - \beta$ , and consequently  $T \cap \partial S'$  is nonempty. But this implies  $T \cap \partial S' = \partial S'$ , since  $\partial S'$  is connected. Since  $T$  was arbitrary, every component of  $S' - \cup \beta_i$  contains  $\partial S'$ , and therefore  $S' - \cup \beta_i$  is connected and  $\cup \beta_i$  is non-separating, as claimed.

It follows that  $S' - \cup_i \beta_i$  is connected, and therefore (by reason of  $\chi$ ) is homeomorphic to a sphere with  $2n+3$  holes. We can therefore find  $n+1$  disjoint properly embedded arcs in  $S' - \cup_i \beta_i$ , each of which runs between the two boundary components corresponding to a single  $\beta_i$ . This proves the claim, and therefore the lemma.  $\square$

As a corollary, we obtain the following inequality:

**Proposition 2.27** ( $[x, w]$  inequality). *For any word  $w$ , the word  $[x, w]$  (where  $x$  is an extra free generator) satisfies*

$$l([x, w]^{2n+1} | [x, w]) \leq n + l(w^{n+1} | w)$$

and consequently

$$\text{sl}([x, w] | [x, w]) \leq \frac{1 + \text{sl}(w | w)}{2}$$

*Proof.* The proposition follows from Lemma 2.26, together with the observation that  $[x, w^{n+1}]$  can be written as a product of  $l(w^{n+1} | w)$   $[x, w]$ -words, by bullet (3) of Lemma 2.2.  $\square$

As an important special case, one deduces:

**Proposition 2.28** ( $\gamma_n$  inequality). *For any  $n$  there is an inequality*

$$\text{sl}(\gamma_n | \gamma_n) \leq 1 - 2^{1-n}$$

*Proof.* For  $n = 1$  this follows from  $\text{sl}(x | x) = 0$ , a special case of Lemma 2.13. Then the result follows from Proposition 2.27 and induction on  $n$ .  $\square$

**Question 2.29.** *Is the estimate in Proposition 2.28 sharp?*

### 3. GEOMETRY OF $G_W$

In this section we begin a more systematic study of stable  $W$ -length, from a geometric point of view. The main tool is the geometry of Cayley graphs, and their asymptotic cones. A verbal subgroup of any group admits a tautological, characteristic, bi-invariant metric, and it is the asymptotic geometry of this metric that gives us insight into stable  $W$ -length.

**3.1. Cayley graph  $C_W$ .** Let  $G$  be a group, and  $G_W$  its verbal subgroup associated to some subset  $W \subset F$ . If  $w$  is a single element, we assume it is reflexive. Otherwise, we assume that  $W$  is symmetric; i.e.  $W = W^{-1}$ .

Let  $C_W(G)$  (or just  $C_W$  or  $C$  if  $G$  or  $W$  are understood) denote the Cayley graph of  $G_W$  with all  $W$ -words in  $G$  as generators. We give  $C$  the structure of a (path) metric space, by declaring that each edge has length 1. This induces a metric on  $G_W$  which we call the  $W$ -metric, where the distance from  $g$  to  $h$  is the  $W$ -length of  $g^{-1}h$ .

**Lemma 3.1.** *The  $W$ -metric is both left and right invariant for the action of  $G_W$  on itself. Furthermore,  $\text{Aut}(G)$  acts on  $G_W$  by isometries, and any homomorphism  $G \rightarrow H$  induces a 1-Lipschitz (simplicial) map from  $C_W(G)$  to  $C_W(H)$ .*

*Proof.* For any  $f, g, h \in G_W$  we have

$$d(fg, fh) = l(g^{-1}h | W) = d(g, h)$$

and

$$d(gf, hf) = l((g^{-1}h)^{f^{-1}} | W) = l(g^{-1}h | W) = d(g, h)$$

The latter two properties are restatements of Lemma 2.14 in geometric language.  $\square$

From the construction, the stable  $W$ -length of an element  $g \in G_W$  is just the *translation length*  $\tau(g)$  of  $g$  on  $G_W$ ; i.e. the limit  $\tau(g) = \lim_{n \rightarrow \infty} d(\text{id}, g^n)/n$ .

The large-scale geometry of commutator subgroups (i.e. the case  $w = [x, y]$ ) was studied in [8].

**3.2. Asymptotic cone.** Though  $C_W$  is an interesting geometric object in its own right, stable  $W$ -length can be more easily studied in an object derived from  $C_W$ , namely the *asymptotic cone*; see e.g. Gromov [13] Chapter 2 (especially p. 36) for an introduction to asymptotic cones in group theory and their properties. The construction of an asymptotic cone depends (in general) on a highly non-constructive choice, namely the choice of a non-principal ultrafilter. However, we will shortly restrict attention to a subset of the asymptotic cone whose geometry does not depend on any choices, and therefore our discussion of such objects is very terse.

Essentially, a non-principal ultrafilter  $\omega$  picks out a limit of every absolutely bounded sequence  $a_n$ , which we denote  $\lim_{\omega} a_n$ . We can think of  $\lim_{\omega}$  as a function from  $\ell^{\infty}$  to  $\mathbb{R}$ ; this function is a ring homomorphism. The “limit”  $\lim_{\omega}$  is always contained between  $\limsup$  and  $\liminf$ , and is equal to the honest limit of some infinite subsequence of the  $a_n$  (so, for instance, if  $\lim_{n \rightarrow \infty} a_n$  exists it is equal to  $\lim_{\omega} a_n$ ) but otherwise satisfies no *a priori* constraint. For more details, see Gromov *op. cit.*

Fix (once and for all) a non-principal ultrafilter  $\omega$ .

**Definition 3.2.** Given a subset  $W \subset F$  and a group  $G$ , define the asymptotic cone  $\hat{A}_W(G)$  (or just  $\hat{A}_W$  or  $\hat{A}$  if  $G$  is understood) to be the ultralimit of the sequence of metric spaces  $C_W(G)$  with the  $W$ -metric rescaled by a factor of  $1/n$ .

Thus a point in  $\hat{A}_W(G)$  is an equivalence class of sequence  $\{a_n\}$  with  $d(\text{id}, a_n) = O(n)$ , where  $d(\{a_n\}, \{b_n\}) = \lim_{\omega} d(a_n, b_n)/n$ , and where  $\{a_n\}$  and  $\{b_n\}$  are in the same equivalence class (denoted  $\{a_n\} \sim \{b_n\}$ ) iff  $\lim_{\omega} d(a_n, b_n)/n = 0$ .

**Proposition 3.3.** *The asymptotic cone  $\hat{A}_W$  has the structure of a group, with a bi-invariant metric, satisfying the law  $W$ .*

*Proof.* The group structure on  $\hat{A}$  is defined by taking  $\{\text{id}\}$  (the constant equivalence class) as the identity, defining multiplication by  $\{a_n\} \cdot \{b_n\} = \{a_n b_n\}$ , and inverses by  $\{a_n\}^{-1} = \{a_n^{-1}\}$ . To see that this is well-defined, suppose  $\{a_n\} \sim \{a'_n\}$  and  $\{b_n\} \sim \{b'_n\}$ . By definition, we have  $a_n = w_n a'_n$  and  $b_n = v_n b'_n$  where  $l(w_n|W) = d(a_n, a'_n)$  and  $l(v_n|W) = d(b_n, b'_n)$ . Hence

$$a_n b_n = w_n a'_n v_n b'_n = w_n v_n^{a'_n} a'_n b'_n$$

so

$$\lim_{\omega} d(a_n b_n, a'_n b'_n)/n \leq \lim_{\omega} (d(a_n, a'_n) + d(b_n, b'_n))/n = 0$$

The bi-invariance of the metric on  $\hat{A}$  follows immediately from the definition of the multiplication, and the bi-invariance of the metric on  $C$ .

Finally, if  $\{a_{1,n}\}\{a_{2,n}\} \cdots \{a_{m,n}\}$  is a  $W$ -word in  $\hat{A}$  (with its group structure), then  $a_{1,n} a_{2,n} \cdots a_{m,n}$  is a  $W$ -word in  $G$  for each  $n$ , and therefore

$$\{a_{1,n}\}\{a_{2,n}\} \cdots \{a_{m,n}\} = \{a_{1,n} a_{2,n} \cdots a_{m,n}\} \sim \{\text{id}\}$$

□

**3.3. The group  $B_W$ .** We now relate  $\hat{A}_W$  to  $G_W$  (in fact, to a slightly larger group depending on  $G$  and  $W$ ).

**Lemma 3.4.** *Let  $W \subset F$  be given, and let  $g_1, g_2, \dots, g_m$  be in  $G_W$ . There is an estimate*

$$l((g_1^{a_1} g_2^{a_2} \cdots g_m^{a_m})^{-1} (g_1^{b_1} g_2^{b_2} \cdots g_m^{b_m}) | W) \leq \sum_i l(g_i^{|b_i - a_i|} | W)$$

*Proof.* By expanding,

$$\begin{aligned} \text{LHS} &= g_m^{-a_m} g_{m-1}^{-a_{m-1}} \cdots g_2^{-a_2} g_1^{b_1 - a_1} g_2^{b_2} \cdots g_m^{b_m} \\ &= g_m^{-a_m} g_{m-1}^{-a_{m-1}} \cdots g_2^{b_2 - a_2} \cdots g_m^{b_m} (g_1^{b_1 - a_1})^* \\ &= \prod_i (g_i^{b_i - a_i})^* \end{aligned}$$

□

*Remark 3.5.* It would be nice to generalize Lemma 3.4 to certain homogeneous expressions  $g_1 g_2 \cdots g_m$  whose *product* is in  $G_W$ , even if the individual  $g_i$  are not. In fact, this can be done if  $m = 2$ , but *not* for higher  $m$  unless  $G_W = [G, G]$  (i.e. unless every commutator is a product of  $W$ -words).

Suppose  $gh \in G_W$ . Then

$$\begin{aligned} g^n h^n &= (gh)^n [g^{n-1}, h]^* [g^{n-2}, h]^* \cdots [g, h]^* \\ &= (gh)^n [g^{n-1}, gh]^* \cdots [g, gh]^* \end{aligned}$$

where we use the identity  $[a, b] = [a, ab]^*$ . Since any commutator  $[*, gh]$  is a product of at most  $2 \cdot l(gh|W)$   $W$ -words, we see that  $l(g^n h^n | W) \leq 3n \cdot l(gh|W)$ . Moreover,

$$(g^{n_1} h^{n_1})^{-1} (g^{n_2} h^{n_2}) = (g^{n_2 - n_1} h^{n_2 - n_1})^*$$

and we derive the inequality

$$l((g^{n_1} h^{n_1})^{-1} (g^{n_2} h^{n_2}) | W) \leq 3|n_2 - n_1| \cdot l(gh|W)$$

which is analogous to the estimate in Lemma 3.4.

On the other hand, in general the condition that  $xyz$  is in  $G_W$  does *not* imply that  $x^n y^n z^n$  is in  $G_W$  for all  $n$  (in fact, for any  $n$  other than 0, 1), unless  $G_W = [G, G]$ . For example, if  $y, z$  are arbitrary and  $x$  is chosen so that  $xyz \in G_W$ , then

$$x^{-1} y^{-1} z^{-1} = x^{-1} z^{-1} y^{-1} [y, z]$$

Now,  $x^{-1} z^{-1} y^{-1}$  is in  $G_W$  because  $xyz$  is, so the left hand side is in  $G_W$  iff  $[y, z]$  is. But  $y$  and  $z$  are arbitrary, so this holds for all  $y, z$  if and only if  $G_W = [G, G]$ , as claimed.

Let  $B_W$  be the free group obeying the law  $W$  generated by the elements of  $G_W$  as a *set*, subject to the relation of homogeneity  $g^n = g^{\star n}$  where  $\star$  is the group operation in  $B_W$ . There is a natural inclusion  $G_W \rightarrow B_W$  of *sets*, taking each element of  $G_W$  to the corresponding generator.

**Lemma 3.6.** *There is a homomorphism  $\rho : B_W \rightarrow \hat{A}_W$  defined by the formula*

$$\rho(g_1 \star g_2 \star \cdots \star g_m) = \{g_1^n g_2^n \cdots g_m^n\}$$

*Proof.* From the definition of the group operation in  $\hat{A}_W$ , the right hand side is equal to the product  $\{g_1^n\}\{g_2^n\}\cdots\{g_m^n\}$ , so the map is a homomorphism. To see that it is well-defined, observe that all relations (i.e.  $W$ -words, and the homogeneity relations) are satisfied in  $\hat{A}_W$  by Proposition 3.3, so the map is well-defined.  $\square$

*Remark 3.7.* It is not typically true that this homomorphism is an injection, even if  $G$  is free.

This homomorphism lets us map  $G_W$  to  $\hat{A}_W$  by  $g \rightarrow \{g^n\}$ . From the definitions,  $\text{sl}(g|W) = d(\{\text{id}\}, \{g^n\})$ ; in other words *stable w-length can be recovered from the geometry of  $\hat{A}_W$  and the map  $B_W \rightarrow \hat{A}_W$ .*

**3.4. Real structure.** For any group  $G$ , let  $B_W \otimes \mathbb{R}$  be the free group obeying the law  $W$  generated by expressions of the form  $g^t$  where  $g \in G_W$  and  $t \in \mathbb{R}$ , subject to the relation of homogeneity  $g^s \star g^t = g^{s+t}$ . We topologize  $B_W \otimes \mathbb{R}$  with the weakest topology for which multiplication and inverse are continuous, as well as each homomorphism  $\mathbb{R} \rightarrow B_W \otimes \mathbb{R}$  of the form  $t \rightarrow g^t$ .

We extend  $\rho : B_W \rightarrow \hat{A}_W$  to  $\rho : B_W \otimes \mathbb{R}$  by defining

$$g^t \rightarrow \{g^{\lfloor tn \rfloor}\}$$

on the  $\mathbb{R}$  subgroups, and extending to arbitrary products by using the group multiplication. Hence

$$g_1^{t_1} g_2^{t_2} \cdots g_m^{t_m} \rightarrow \{g_1^{\lfloor t_1 n \rfloor} g_2^{\lfloor t_2 n \rfloor} \cdots g_m^{\lfloor t_m n \rfloor}\}$$

Notice that although  $g^{\lfloor tn \rfloor}$  is not necessarily the inverse of  $g^{\lfloor -tn \rfloor}$  if  $tn$  is not an integer, nevertheless these elements are distance at most  $2l(g|W)$  apart in  $C_W$ , and therefore  $\{g^{\lfloor tn \rfloor}\}$  and  $\{g^{\lfloor -tn \rfloor}\}$  are inverse in  $\hat{A}_W$ .

**Definition 3.8.** The *real cone*  $A_W$  is the image of  $B_W \otimes \mathbb{R}$  in  $\hat{A}_W$ .

For the remainder of the paper we restrict attention to  $A_W$ . The metric on  $\hat{A}_W$  restricts to a (bi-invariant) metric on  $A_W$ , consequently giving it the structure of a topological group. In the sequel, we use the notation  $\|g\|$  for  $d(\text{id}, g)$ , where  $g \in A_W$ .

The real structure on  $B_W \otimes \mathbb{R}$  gives rise to a natural (multiplicative)  $\mathbb{R}$  action on  $A_W$ .

**Lemma 3.9.** *There is a continuous family of endomorphisms  $\mathbb{R} \times A_W \rightarrow A_W$  with the following properties:*

- (1) *the action is multiplicative — i.e.  $1 \times$  is the identity on  $A_W$ , and  $\lambda \times (\mu \times g) = (\lambda\mu) \times g$  for  $\lambda, \mu \in \mathbb{R}$  and  $g \in A_W$ ; and*
- (2) *for any  $g \in A_W$  and  $\lambda \in \mathbb{R}^*$ , we have  $\|\lambda \times g\| = |\lambda| \cdot \|g\|$*

*Proof.* Let  $B$  be the group freely generated by expressions of the form  $g^t$  with  $g \in G$  and  $t \in \mathbb{R}$ , subject to the relations  $g^t g^s = g^{t+s}$ . There is a natural  $\mathbb{R}$  action on  $G$ , given by  $\lambda \times g^t = g^{\lambda t}$ . This action evidently preserves the  $W$ -subgroup of  $B$ , and therefore descends to an automorphism of  $B_W \otimes \mathbb{R}$ . We claim that this action preserves the kernel of  $\rho$ , and thereby defines an action on  $A_W$ . In other words, we

need to show that if  $\|\rho(g_1^{t_1} \cdots g_m^{t_m})\| = 0$  then  $\|\rho(g_1^{\lambda t_1} \cdots g_m^{\lambda t_m})\| = 0$  for any  $\lambda$ . But

$$\begin{aligned} \|\rho(g_1^{\lambda t_1} \cdots g_m^{\lambda t_m})\| &= d(\{\text{id}\}, \{g_1^{\lfloor \lambda t_1 n \rfloor} \cdots g_m^{\lfloor \lambda t_m n \rfloor}\}) \\ &= \lim_{\omega} \frac{1}{n} l(g_1^{\lfloor \lambda t_1 n \rfloor} \cdots g_m^{\lfloor \lambda t_m n \rfloor} | W) \\ &= \lambda \cdot \lim_{\omega} \frac{1}{n} l(g_1^{\lfloor t_1 n \rfloor} \cdots g_m^{\lfloor t_m n \rfloor} | W) \\ &= \lambda \cdot \|\rho(g_1^{t_1} \cdots g_m^{t_m})\| \end{aligned}$$

(where the third line follows from the second by an approximate change of variables and elementary estimates) thereby establishing both claims.  $\square$

The existence of this  $\mathbb{R}$  action has the following topological consequence.

**Lemma 3.10.**  *$A_W$  is contractible and locally contractible.*

*Proof.* The  $\mathbb{R}$  action defines a deformation retraction of  $A_W$  to the identity element, where each element  $g$  moves along the path  $(1 - t) \times g$  where  $t$  goes from 0 to 1. This retraction takes the ball of radius  $r$  around the identity inside itself for all positive  $t$ , so  $A_W$  is locally contractible.  $\square$

Since  $\hat{A}_W$  obeys the law  $W$ , so does  $A_W$ . However, under certain circumstances we can say much more about  $A_W$ . The main theorem of this section is the following:

**Theorem 3.11.** *Suppose  $A_W$  is nilpotent. Then it is a normed vector space (in particular it is abelian).*

*Proof.* If  $A_W$  is abelian, then the existence of the  $\mathbb{R}$  action with the properties proved in Lemma 3.9 shows that  $A_W$  is a normed vector space, with norm  $\|\cdot\|$ . So it suffices to show  $A_W$  is abelian.

For legibility, denote the  $n$ th element of the lower central series of  $A_W$  by  $A_n$ , so that  $A_1 = A_W$  and  $A_n = 0$  for some  $n$ . Suppose we have shown for some integer  $k$ , for all real  $t$ , for all  $h \in A_k$  and all  $g \in G_W$ , that the commutator  $[\rho(g^t), h] = \text{id}$ . Since  $A_W$  is generated by elements of the form  $\rho(g^t)$  this implies that  $A_{k+1} = 0$ .

Let  $h \in A_{k-1}$ . Then

$$\begin{aligned} [\rho(g^t), h] &= [\rho(g^{t/2}), h][\rho(g^{t/2}), h][[h, \rho(g^{t/2})], \rho(g^{t/2})^h] && \text{by Lemma 2.2} \\ &= [\rho(g^{t/2}), h][\rho(g^{t/2}), h] && \text{because } A_{k+1} = 0 \\ &= [\rho(g^{t/2}), h^2][[h, \rho(g^{t/2})], h]^{-1} && \text{by Lemma 2.2} \\ &= [\rho(g^{t/2}), h^2] && \text{because } A_{k+1} = 0 \end{aligned}$$

By induction,  $[\rho(g^t), h] = [\rho(g^{t/2^m}), h^{2^m}]$  for every positive integer  $m$ . Hence

$$\|[\rho(g^t), h]\| = \|[\rho(g^{t/2^m}), h^{2^m}]\| = \|g^{t/2^m} \cdot (g^{-t/2^m})^*\| \leq 2 \cdot 2^{-m} \|g^t\|$$

Since  $m$  is arbitrary,  $[\rho(g^t), h] = \text{id}$  and therefore (assuming  $k \geq 2$ ),  $A_k = 0$ . Since  $A_W$  is nilpotent by hypothesis,  $A_n = 0$  for some  $n$ , and therefore  $A_W$  is abelian, as claimed.  $\square$

As a corollary, one obtains the following propositions.

**Proposition 3.12.** *Suppose  $A_W$  is locally compact. Then it is a (finite dimensional) normed vector space.*

*Proof.* Since  $A_W$  is connected, locally path connected (by Lemma 3.10) and locally compact by hypothesis, the Gleason-Montgomery-Zippin theorem (i.e. the affirmative solution of Hilbert's fifth problem [11, 16]) shows that  $A_W$  is a Lie group. Since it obeys a law, it is necessarily solvable, since non-solvable Lie groups contain nonabelian free subgroups.

By Theorem 3.11, it suffices to show it is nilpotent. Equivalently, we need to show that for each  $g \in G_W$  and  $t \in \mathbb{R}$  the conjugation action of  $\rho(g^t)$  on the derived subgroup  $A'$  is trivial.

Suppose not, so that  $\rho(g^t)$  is a 1-parameter family of isometric automorphisms of  $A'$ . Since  $A'$  is finite-dimensional, the group of isometric automorphisms is compact. In particular, there are arbitrarily large values of  $t$  for which  $\rho(g^t)$  is arbitrarily close to the identity in the isometry group of  $A'$ .

By the compactness of the group of isometric automorphisms of  $A'$  there is a constant  $C$  so that if  $\iota$  is an automorphism with  $d(h, \iota(h)) < \epsilon$  for all  $h$  in the ball about the identity in  $A'$  of radius  $1/2$ , then  $d(h, \iota(h)) < C \cdot \epsilon$  for all  $h$  in the ball about the identity in  $A'$  of radius 1 (in fact, since the metric on  $A'$  is bi-invariant, we can take  $C = 2$ , but this is unnecessary for our argument).

Now, for any positive  $\epsilon$ , there are arbitrarily large  $t$  such that  $d(h, h^{\rho(g^t)}) < \epsilon$  for all  $h$  in the ball of radius 1 in  $A'$ . Consequently, for all  $s \in [1/2, 1]$  and all  $h$  in the ball of radius 1 in  $A'$ ,

$$d(s \times h, s \times (h^{\rho(g^t)})) = d(s \times h, (s \times h)^{\rho(g^{ts})}) < s\epsilon$$

It follows that  $d(h, h^{\rho(g^{ts})}) < s\epsilon C$  for all  $h$  in the ball of radius 1 in  $A'$ , and all  $s \in [1/2, 1]$ . It follows that  $d(h, h^{\rho(g^{t'})}) < 2\epsilon C$  for all  $t' \in [0, t/2]$ .

Since  $\epsilon$  is arbitrarily small, since  $C$  is fixed, and since  $t$  can be chosen arbitrarily large for each  $\epsilon$ , it follows that the conjugation action of  $\rho(g^t)$  is trivial for all  $g \in G_W$  and  $t \in \mathbb{R}$ . Hence  $A'$  is central, and therefore  $A_W$  is nilpotent, and therefore abelian and a finite dimensional normed vector space, as claimed.  $\square$

*Remark 3.13.* In fact, we do not really need the full power of Gleason-Montgomery-Zippin. We only need to know that the group of isometries of  $A'$  is compact (which follows from Arzela-Ascoli and local compactness) to deduce that  $A$  is abelian; then one may appeal to Pontriagin's solution of Hilbert's fifth problem for locally compact abelian groups.

Recall the notation  $\beta_2 := [[x, y], [z, w]]$  for free generators  $x, y, z, w$ .

**Theorem 3.14.** *Suppose in some group  $G$  that stable commutator length vanishes identically. Then  $A_W$  is abelian (and hence a normed vector space) for  $W = \beta_2$ .*

*Proof.* By Theorem 3.11 it suffices to show that  $A_W$  is nilpotent. We show it satisfies the law  $[x, [y, z]] = 0$ .

Let  $f, g, h$  be arbitrary elements of  $A_W$ , and suppose  $f = \rho(f_1^{t_1} f_2^{t_2} \cdots f_m^{t_m})$ . We write  $g = \{g_n\}$ ,  $h = \{h_n\}$  and  $f = \{f_1^{\lfloor nt_1 \rfloor} f_2^{\lfloor nt_2 \rfloor} \cdots f_m^{\lfloor nt_m \rfloor}\}$ . By hypothesis, the commutator length of each expression  $f_i^{\lfloor nt_i \rfloor}$  is  $o(n)$ . It follows that

$$[f, [g, h]] = \{[f_1^{\lfloor nt_1 \rfloor} f_2^{\lfloor nt_2 \rfloor} \cdots f_m^{\lfloor nt_m \rfloor}, [g_n, h_n]]\} = \{o(n) \beta_2 - \text{words}\} = 0$$

where we use Lemma 2.2 bullet (4) in the second step. Hence  $A_W$  is nilpotent and therefore abelian.  $\square$

*Remark 3.15.* Recall that  $\beta_1 = [x, y]$ , so that  $\beta_1$ -length is just commutator length. A similar argument shows that if stable  $\beta_n$ -length vanishes identically in  $G$ , then for  $W = \beta_{n+1}$  the group  $A_W$  obeys the law  $[x, \beta_n]$ .

**3.5.  $\gamma_n$ -quasimorphisms.** For the remainder of this section we typically specialize to the case that  $w = \gamma_n$  (recall  $\gamma_2 = [x, y]$ ,  $\gamma_3 = [x, [y, z]]$  and so on). Inspired by the phenomenon of (generalized) *Bavard duality* in the theory of stable commutator length, we make the following definition:

**Definition 3.16.** A *homogeneous  $W$ -quasimorphism* (hereafter  *$W$ -hoq*) is a Lipschitz homomorphism from  $A_W$  to  $\mathbb{R}$ .

If  $\phi$  is a  $W$ -hoq, let  $d(\phi)$  denote the optimal Lipschitz constant.

For general  $W$ , it is not clear whether any homomorphisms from  $A_W$  to  $\mathbb{R}$  exist, let alone Lipschitz ones. But when  $A_W$  is a normed vector space (which holds, for instance, if  $W = \gamma_n$ , by Theorem 3.11) the Hahn-Banach theorem guarantees the existence of a rich supply of  $W$ -hoqs.

In fact, we have the following proposition:

**Lemma 3.17.** *For any  $n$  and any  $g \in G_n$ , there is an equality*

$$\text{sl}(g|\gamma_n) = \sup_{\phi} \phi(\rho(g))/d(\phi)$$

where the supremum is taken over all  $\gamma_n$ -hoqs  $\phi$ .

*Proof.* This follows from the fact that  $\|\rho(g)\| = \text{sl}(g|\gamma_n)$ , together with Theorem 3.11 and the Hahn-Banach theorem.  $\square$

Whenever  $A_W$  is a vector space, it is spanned by the vectors  $\rho(g)$  for  $g \in G_W$ . Since a  $W$ -hoq  $\phi$  is a homomorphism, such a function on  $A_W$  is determined by its values on  $\rho(g)$ . Therefore, by abuse of notation, we think of a  $W$ -hoq in this case as a function on  $G_W$ ; thus the equality from proposition 3.17 would be expressed in the form  $\text{sl}(g|\gamma_n) = \sup_{\phi} \phi(g)/d(\phi)$ . On the other hand, it is not evident from the definition how to recover (or to estimate)  $d(\phi)$  directly from the values of  $\phi$  on the elements of  $G_W$ . We now address this issue.

**Definition 3.18.** A *weak  $\gamma_n$ -hoq* is a function  $\phi : G \rightarrow \mathbb{R}$  for which there is a least non-negative real number  $D(\phi)$  (called the *defect*) satisfying the following properties:

- (1) (homogeneity) for any  $g \in G$  and any  $n \in \mathbb{Z}$  there is an equality  $\phi(g^n) = n\phi(g)$ ;
- (2) (quasimorphism) for any  $g, h \in G$  there is an inequality

$$|\phi(g) + \phi(h) - \phi(gh)| \leq D(\phi) \min(l(g|\gamma_{n-1}), l(h|\gamma_{n-1}), l(gh|\gamma_{n-1}))$$

*Remark 3.19.* For  $n = 2$ , this is precisely the classical definition of a homogeneous quasimorphism. For,  $\gamma_1 = x$  and therefore  $l(g|\gamma_1) = 1$  if  $g \neq \text{id}$  and 0 otherwise.

From the definition one can deduce some basic properties of weak  $\gamma_n$ -hoqs.

**Lemma 3.20.** *Suppose  $G$  is perfect. Then any weak  $\gamma_n$ -hoq  $\phi : G \rightarrow \mathbb{R}$  satisfies the following properties:*

- (1)  $\phi$  is a class function;
- (2) if  $h$  is a  $\gamma_n$ -word, then  $\phi(h) \leq D(\phi)$ ;



(3) for any  $g \in G$  there is an estimate

$$\phi(g) \leq (2l(g|\gamma_n) - 1)D(\phi)$$

and consequently

$$sl(g|\gamma_n) \geq \phi(g)/2D(\phi)$$

*Proof.* The hypothesis that  $G$  is perfect implies (inductively) that  $l(g|\gamma_m)$  is finite for every  $g \in G$  and every integer  $m$ . To see that  $\phi$  is a class function, observe that for any elements  $g, h \in G$  and any  $n$  one has

$$|\phi(hg^n h^{-1}) - \phi(h) - \phi(g^n) - \phi(h^{-1})| \leq 2D(\phi)l(g|\gamma_{n-1})$$

By homogeneity, the left hand side is equal to  $n \cdot |\phi(hgh^{-1}) - \phi(g)|$  whereas the right hand side is a constant independent of  $n$ , thus proving the first claim.

Now suppose  $h$  is a  $\gamma_n$ -word; i.e.  $h = [x, y]$  for some  $x$  and  $y$  where  $y$  is a  $\gamma_{n-1}$ -word. Then

$$|\phi(h) - \phi(xy x^{-1}) - \phi(y^{-1})| \leq D(\phi)l(y|\gamma_{n-1}) = D(\phi)$$

On the other hand, since  $\phi$  is a class function and homogeneous, the left hand side is equal to  $|\phi(h)|$  and the inequality is proved.

Finally, if we express  $g = h_1 h_n \cdots h_m$  where each  $h_i$  is a  $\gamma_n$ -word, then by induction  $\phi(g) \leq \sum \phi(h_i) + (m-1)D(\phi) \leq (2m-1)D(\phi)$ . This proves bullet (3).  $\square$

**Proposition 3.21.** *The set of weak  $\gamma_n$ -hoqs on  $G$  is a real vector space, which we denote  ${}_nQ(G)$ , and  $D(\cdot)$  is a semi-norm. If  $G$  is perfect,  ${}_nQ(G)$  is a Banach space with the norm  $D(\cdot)$ .*

*Proof.* The defining properties of a weak  $\gamma_n$ -hoq are an (infinite) system of linear equalities and inequalities. It follows by inspection that  ${}_nQ(G)$  is a real vector space, and that  $D(\cdot)$  is a semi-norm.

If  $G$  is perfect, then  $l(g|\gamma_{n-1})$  is finite for any  $g$ , and therefore  $D(\phi) = 0$  if and only if  $\phi$  is a homomorphism to  $\mathbb{R}$ , which is necessarily trivial for  $G$  perfect. Thus  $D(\cdot)$  is a norm, and it remains to show that  ${}_nQ(G)$  is complete in this norm.

So, let  $\phi_i$  be a sequence of weak  $\gamma_n$ -hoqs for which  $D(\phi_i - \phi_j)$  is a Cauchy sequence; i.e. for all  $\epsilon > 0$  there is an  $N$  so that  $D(\phi_i - \phi_j) < \epsilon$  for  $i, j > N$ . By bullet (3) from Lemma 3.20 it follows that the values of the  $\phi_i$  on any element  $g$  form a Cauchy sequence, and therefore the  $\phi_i$  converge pointwise to a limit  $\phi$ . A pointwise limit of homogeneous functions is homogeneous; moreover, for any  $g, h$  pointwise convergence implies

$$|(\phi - \phi_j)(g + h - gh)| \leq \limsup_i D(\phi_i - \phi_j) \min(l(g|\gamma_{n-1}), l(h|\gamma_{n-1}), l(gh|\gamma_{n-1}))$$

and therefore  $\phi$  is in  ${}_nQ(G)$  and  $D(\phi_i - \phi) \rightarrow 0$ , and the lemma is thereby proved.  $\square$

The following lemma justifies the terminology “weak  $\gamma_n$ -hoq”.

**Lemma 3.22.** *If  $G$  is perfect, every  $\gamma_n$ -hoq is a weak  $\gamma_n$ -hoq with  $D(\phi) \leq d(\phi)$ .*

*Proof.* Let  $\phi : A_{\gamma_n} \rightarrow \mathbb{R}$  be a  $d(\phi)$ -Lipschitz homomorphism, and by abuse of notation we let  $\phi : G \rightarrow \mathbb{R}$  be the function defined by  $\phi(g) := \phi(\rho(g))$ . We estimate

$$\begin{aligned}
|\phi(g) + \phi(h) - \phi(gh)| &= |\phi(\rho(g * h * (h^{-1}g^{-1})))| \\
&\leq d(\phi) \|\{g^m h^m h^{-1}g^{-1} \dots h^{-1}g^{-1}\}\| \\
&= d(\phi) \lim_{m \rightarrow \infty} \frac{1}{m} l(g^m h^m h^{-1}g^{-1} \dots h^{-1}g^{-1} | \gamma_n) \\
&\leq d(\phi) \lim_{m \rightarrow \infty} \frac{1}{m} l([g, h]^* [g^2, h]^* \dots [g^{m-1}, h]^* | \gamma_n) \\
&\leq d(\phi) l(h | \gamma_{n-1})
\end{aligned}$$

By symmetry, we obtain the estimate

$$|\phi(g) + \phi(h) - \phi(gh)| \leq d(\phi) \min(l(g | \gamma_{n-1}), l(h | \gamma_{n-1}), l(gh | \gamma_{n-1}))$$

so that  $\phi$  is a weak  $\gamma_n$ -hoq with  $D(\phi) \leq d(\phi)$ , as claimed.  $\square$

We therefore obtain the following generalization of Bavard duality.

**Theorem 3.23** ( $\gamma_n$ -Duality theorem). *For any perfect group  $G$ , and any  $g \in G$  there is an inequality*

$$\sup_{\phi} \phi(g)/D(\phi) \geq \text{sl}(g | \gamma_n) \geq \sup_{\phi} \phi(g)/2D(\phi)$$

where the supremum is taken over all weak  $\gamma_n$ -hoqs.

*Proof.* The upper bound follows from Lemma 3.22 and Lemma 3.17, and the lower bound from Lemma 3.20.  $\square$

An interesting application of this duality theorem is an *a priori* estimate of the ratio of  $\text{scl}(\ast)$  and  $\text{sl}(\ast | \gamma_n)$  in any perfect group. This estimate follows from the surprising fact that one may bound from below the stable  $\gamma_{n-1}$  length of a word from the value of a weak  $\gamma_n$ -hoq, providing  $n \geq 3$ .

**Lemma 3.24.** *Let  $G$  be a perfect group. Then any weak  $\gamma_n$ -hoq  $\phi : G \rightarrow \mathbb{R}$  satisfies the following properties:*

- (1) *if  $h$  is a  $\gamma_{n-1}$ -word, then  $\phi(h) \leq D(\phi)$*
- (2) *for any  $g \in G$  there is an estimate*

$$\phi(g) \leq (2l(g | \gamma_{n-1}) - 1)D(\phi)$$

and consequently

$$\text{sl}(g | \gamma_{n-1}) \geq \phi(g)/2D(\phi)$$

*Proof.* Let  $g$  be arbitrary, and let  $k$  be a  $\gamma_{n-2}$ -word. There is a constant  $C$  so that for any  $m$  we have  $l(g^m k^m (gk)^{-m} | \gamma_n) \leq Cm$  and therefore by Lemma 3.20 we can estimate

$$|\phi(g^m k^m (gk)^{-m})| \leq C'm$$

for some constant  $C'$  depending only on  $g$ ,  $k$  and  $D(\phi)$ .

On the other hand, we can write

$$\begin{aligned}
g^m k^m (gk)^{-m} &= g^m k^m k^{-1} g^{-1} k^{-1} g^{-1} \dots k^{-1} g^{-1} \\
&= g^m k^{m-1} g^{-1} k^{-1} \dots k^{-1} g^{-1} \\
&= g^m k^{m-2} g^{-1} [g, k] g^{-1} k^{-1} \dots k^{-1} g^{-1} \\
&= g^m k^{m-2} g^{-1} g^{-1} k^{-1} \dots k^{-1} g^{-1} [g, k]^* \\
&= g^m k^{m-3} g^{-1} g^{-1} g^{-1} k^{-1} \dots k^{-1} g^{-1} [g, k]^* [g, k]^* [g, k]^* \\
&= \text{product of } m(m-1)/2 \text{ conjugates of } [g, k]
\end{aligned}$$

By abuse of notation, for any  $r$  we denote a product of  $r$  conjugates of  $[g, h]$  by  $[g, h]^{*r}$ . Hence we have  $g^m k^m (gk)^{-m} = [g, k]^{*m(m-1)/2}$ . On the other hand, for any  $r$ ,

$$|\phi([g, k]^{*r}) - \phi([g, k]^{*r-1}) - \phi([g, k])| \leq D(\phi)$$

by the defining property of an  $\gamma_n$ -hoq, since  $\phi$  is a class function, and  $l(k|\gamma_{n-2}) = 1$  implies  $l([g, k]|\gamma_{n-1}) \leq 1$ . By induction and the triangle inequality we obtain

$$|\phi([g, k]^{*r}) - r\phi([g, k])| \leq rD(\phi)$$

Hence we conclude

$$|\phi(g^m k^m (gk)^{-m}) - (m(m-1)/2)\phi([g, k])| \leq (m(m-1)/2)D(\phi)$$

Since  $|\phi(g^m k^m (gk)^{-m})| \leq C'm$ , dividing by  $m(m-1)/2$  and taking  $m \rightarrow \infty$  gives the estimate

$$|\phi([g, k])| \leq D(\phi)$$

Since  $g$  was arbitrary, and  $k$  is an arbitrary  $\gamma_{n-2}$ -word, we have proved bullet (1). Bullet (2) follows exactly as in Lemma 3.20: if  $g = h_1 h_2 \dots h_m$  where each  $h_i$  is a  $\gamma_{n-1}$ -word, then by induction

$$\phi(g) \leq \sum \phi(h_i) + (m-1)D(\phi) \leq (2m-1)D(\phi)$$

This completes the proof of the lemma.  $\square$

From this easily follows our Comparison Theorem:

**Theorem 3.25** (Comparison theorem). *For any perfect group  $G$ , for any  $n \geq 2$  and for any  $g \in G$  there is an inequality*

$$2^{n-2} \text{scl}(g) \geq \text{sl}(g|\gamma_n) \geq \text{scl}(g)$$

*Proof.* From Lemma 3.24 and Theorem 3.23 we estimate

$$2 \text{sl}(g|\gamma_{n-1}) \geq \sup_{\phi} \phi(g)/D(\phi) \geq \text{sl}(g|\gamma_n)$$

where the supremum is taken over all weak  $\gamma_n$ -hoqs  $\phi$ . By induction on  $n$ , we get

$$2^{n-2} \text{scl}(g) \geq \text{sl}(g|\gamma_n)$$

The second inequality in the theorem is trivial, since every  $\gamma_n$ -word is a  $\gamma_2$ -word.  $\square$

It follows that weak  $\gamma_n$ -hoqs can be used to estimate scl:

**Corollary 3.26.** *Let  $G$  be a perfect group, and  $\phi$  a weak  $\gamma_n$ -hoq. Then for any  $g \in G$  we have*

$$\text{scl}(g) \geq \phi(g)/2^{n-1}D(\phi)$$

Stable commutator length has many important applications to geometry, topology, dynamics, etc. In general, estimating scl is very difficult; lower bounds are usually obtained from (ordinary homogeneous) quasimorphisms (equivalently, weak  $\gamma_n$ -hoqs when  $n = 2$ ), and by now many interesting constructions of such quasimorphisms are known; see e.g. [6], especially Chapters 3 and 5. In view of Corollary 3.26, it becomes interesting to ask whether there are any natural constructions of weak  $\gamma_n$ -hoqs with  $n > 2$ , arising from geometry or dynamics.

The statement of the Comparison Theorem is purely algebraic, and it is therefore natural to try to find a purely algebraic proof, bypassing the construction of asymptotic cones, the use of Hahn-Banach, etc. In fact, it is not too hard to translate the geometric argument into an algebraic one, though the geometric argument has its own charm. The case of  $\gamma_3$  is especially straightforward; with the same amount of work, one proves a slightly stronger statement.

First, in any group  $G$ , let  $\Gamma_3$  denote the set of words of the form  $[x, y]$  where  $y \in [G, G]$ . The  $\Gamma_3$  words generate the subgroup  $G_3$  (as do the  $\gamma_3$  words, which are special examples of  $\Gamma_3$  words), but there is no uniform comparison between  $\gamma_3$  length and  $\Gamma_3$  length. If  $G$  is perfect, then  $G = [G, G] = G_2$ , so for such groups,  $\Gamma_3$  words are nothing other than commutators, and  $l(*|\Gamma_3) = \text{cl}(*)$ .

**Proposition 3.27.** *In any group  $G$ , for any  $g \in G_3$  there is an inequality*

$$2\text{sl}(g|\Gamma_3) \geq \text{sl}(g|\gamma_3) \geq \text{sl}(g|\Gamma_3)$$

*Proof.* Recall (from Notation 2.12) the notation  $\gamma_3^{\{k\}}$  for an arbitrary product of  $\gamma_3$ -words.

Express  $g^k$  as a product of commutators

$$g^k = [a_1, b_1][a_2, b_2] \cdots [a_m, b_m]$$

where the  $b_i$  are all in  $[G, G]$ , and where  $m/k$  is as close to  $\text{sl}(g|\Gamma_3)$  as we like. Then we have

$$g^{2^n k} = ([a_1, b_1] \cdots [a_m, b_m])^{2^n} = [a_1, b_1]^{2^n} \cdots [a_m, b_m]^{2^n} \gamma_3^{\{m2^n\}}$$

Now, for each  $i$  we have  $[a_i, b_i]^2 = [a_i^2, b_i] \gamma_3^{\{1\}}$  and therefore

$$[a_i, b_i]^{2^n} = [a_i^2, b_i]^{2^{n-1}} \gamma_3^{\{2^{n-1}\}} = [a_i^{2^n}, b_i] \gamma_3^{\{2^n - 1\}}$$

and therefore we can write

$$g^{2^n k} = [a_1^{2^n}, b_1][a_2^{2^n}, b_2] \cdots [a_m^{2^n}, b_m] \gamma_3^{\{m2^n - m\}}$$

We can estimate  $l([a_1^{2^n}, b_1][a_2^{2^n}, b_2] \cdots [a_m^{2^n}, b_m]|\gamma_3) \leq \sum_i l(b_i|\gamma_2)$ , which is a constant independent of  $n$ . Since  $m/k$  is as close as we like to  $\text{scl}(g)$ , and  $n$  is arbitrary, we deduce  $\text{sl}(g|\gamma_3) \leq 2\text{scl}(g)$  as claimed.  $\square$

We do not know whether there is an *a priori* comparison between  $\text{sl}(*|\gamma_3)$  and  $\text{scl}(*)$  on elements of  $G_3$  for an arbitrary group.

**3.6.  $\beta_2$ -quasimorphisms.** By Theorem 3.14, if  $G$  is a perfect group in which scl vanishes identically, then  $A_{\beta_2}$  is a vector space, and  $\text{sl}(g|\beta_2) = \sup_{\phi} \phi(g)/d(\phi)$  where the supremum is taken over all  $\beta_2$ -hoqs  $\phi$ . We would like to obtain a “weak” characterization of  $\beta_2$ -hoqs on such groups, analogous to the definition of weak  $\gamma_n$ -hoqs.

**Definition 3.28.** A *weak  $\beta_2$ -hoq* is a function  $\phi : G \rightarrow \mathbb{R}$  for which there is a least non-negative real number  $D(\phi)$  (called the *defect*) satisfying the following properties:

- (1) (homogeneity) for any  $g \in G$  and any  $n \in \mathbb{Z}$  there is an equality  $\phi(g^n) = n\phi(g)$
- (2) (quasimorphism) for any  $g, h \in G$  there is an inequality

$$|\phi(g) + \phi(h) - \phi(gh)| \leq D(\phi) \min(l(g|\beta_2), l(h|\beta_2), l(gh|\beta_2))$$

The following is the analogue of Lemma 3.22:

**Lemma 3.29.** *If  $G$  is perfect and  $\text{scl}$  vanishes identically, every  $\beta_2$ -hoq is a weak  $\beta_2$ -hoq with  $D(\phi) \leq 2d(\phi)$ .*

*Proof.* As in Lemma 3.22, we must estimate the  $\beta_2$ -length of  $g^m h^m (gh)^{-m}$ . As before, we have an identity

$$g^m h^m (gh)^{-m} = [g, h]^* [g^2, h]^* \cdots [g^{m-1}, h]^*$$

We can write  $h = [a_1, b_1] \cdots [a_r, b_r]$  where each  $a_i, b_i$  is a commutator, and  $r = l(h|\beta_2)$ . Then

$$[g^j, h] = [g^j, [a_1, b_1]]^* [g^j, [a_2, b_2]]^* \cdots [g^j, [a_r, b_r]]^*$$

By the Hall-Witt identity (i.e. bullet (5) from Lemma 2.2) we can write

$$[g^j, [a_i, b_i]] = [a_i^*, [b_i^*, (g^j)^*]] [b_i^*, [a_i^*, (g^j)^*]]$$

and since  $a_i, b_i$  are both commutators by hypothesis, the right hand side is a product of two  $\beta_2$ -words. Hence

$$l(g^m h^m (gh)^{-m} | \beta_2) \leq 2ml(h|\beta_2)$$

By symmetry, the lemma follows.  $\square$

The following is the analogue of Lemma 3.20, though the proof is more circuitous, and has something in common with that of Lemma 3.24:

**Lemma 3.30.** *Suppose  $G$  is perfect, and  $\text{scl}$  vanishes identically in  $G$ . Then any weak  $\beta_2$ -hoq  $\phi : G \rightarrow \mathbb{R}$  satisfies the following properties:*

- (1)  $\phi$  is a class function;
- (2) if  $h$  is a  $\beta_2$ -word, then  $\phi(h) \leq D(\phi)$ ;
- (3) for any  $g \in G$  there is an estimate

$$\phi(g) \leq (2l(g|\beta_2) - 1)D(\phi)$$

and consequently

$$\text{sl}(g|\beta_2) \geq \phi(g)/2D(\phi)$$

*Proof.* To see that  $\phi$  is a class function, observe as in the proof of Lemma 3.20 that for any  $g, h$  and any  $n$ ,

$$|\phi(hg^n h^{-1}) - \phi(h) - \phi(g^n) - \phi(h^{-1})| \leq 2D(\phi)l(h|\beta_2)$$

By homogeneity, and the fact that the right hand side is constant independent of  $n$ , we see that  $\phi(hgh^{-1}) = \phi(g)$ ; i.e.  $\phi$  is a class function.

Secondly, observe that if  $l(g|\beta_2) = 1$  then  $\phi([g, h]) \leq D(\phi)$ . This is because

$$|\phi(hgh^{-1}h^{-1}) - \phi(g) - \phi(hg^{-1}h^{-1})| \leq D(\phi)l(g|\beta_2) = D(\phi)$$

and because  $\phi$  is a homogeneous class function. If we let  $\alpha$  denote the word  $[[u, v], [[x, y], [z, w]]]$  then  $\alpha$  is both a  $\beta_2$ -word, and the commutator of something with a  $\beta_2$  word. Thus by induction,  $|\phi(g)| \leq 2D(\phi)l(g|\alpha)$ . Since  $G$  is perfect, every element has a finite  $\alpha$ -length, so for any  $g$  and  $h$  and any integer  $m$  we have  $|\phi(g^m h^m (gh)^{-m})| \leq Cm$  for some constant  $C$  depending only on  $g, h$  and  $D(\phi)$ .

On the other hand, as in the proof of Lemma 3.24, we can write  $g^m h^m (gh)^{-m}$  as a product of  $m(m-1)/2$  conjugates of  $[g, h]$ . If  $g$  and  $h$  are both commutators,  $[g, h]$  is a  $\beta_2$ -word, and therefore

$$|\phi(g^m h^m (gh)^{-m}) - (m(m-1)/2)\phi([g, h])| \leq (m(m-1)/2)D(\phi)$$

and consequently  $|\phi([g, h])| \leq D(\phi)$ . Since  $g$  and  $h$  are arbitrary commutators, this shows that  $|\phi(g)| \leq D(\phi)$  for any  $\beta_2$ -word  $g$ , proving the second claim.

The third claim follows immediately from this, as in the proof of Lemma 3.20.  $\square$

**Theorem 3.31** ( $\beta_2$ -Duality theorem). *For any perfect group  $G$  in which scl vanishes identically, and for any  $g \in G$  there is an inequality*

$$\sup_{\phi} 2\phi(g)/D(\phi) \geq \text{sl}(g|\beta_2) \geq \sup_{\phi} \phi(g)/2D(\phi)$$

where the supremum is taken over all weak  $\beta_2$ -hoqs.

*Proof.* The upper bound follows from Lemma 3.29, and the lower bound from Lemma 3.30.  $\square$

We conclude this section by making a curious observation on the relation between  $\text{sl}(*|\beta_2)$  and  $l(*|\gamma_3)$ , under the hypothesis that scl vanishes identically.

**Proposition 3.32.** *Suppose  $G$  is perfect, and scl vanishes identically. If  $l(g|\gamma_3) = 1$  then  $\text{sl}(g|\beta_2) \leq 1$ .*

*Proof.* Since for any  $a, b, c$  we have  $[a, b][a, c] = [a, bc][[a, c], b]^*$  by bullet (3) of Lemma 2.2, it follows that

$$[x, [y, z]][x, [y, z]^n] = [x, [y, z]^{n+1}][[x, [y, z]^n], [y, z]]^*$$

for any  $n$  and any  $x, y, z$ . Since  $[[x, [y, z]^n], [y, z]]$  is a  $\beta_2$ -word, it follows by induction that  $[x, [y, z]^n]$  can be written as a product of  $(n-1)$   $\beta_2$ -words with  $[x, [y, z]^n]$ . If scl vanishes identically,  $[y, z]^n$  can be written as a product of  $o(n)$  commutators. If  $G$  is perfect,  $x$  is a product of a finite number of commutators. Hence the  $\beta_2$ -length of  $[x, [y, z]^n]$  is  $o(n)$ , and therefore  $\text{sl}([x, [y, z]]|\beta_2) \leq 1$  as claimed.  $\square$

It seems hard to generalize Proposition 3.32 to estimate  $\text{sl}(*|\beta_2)$  from  $l(*|\gamma_3)$ .

**3.7. Perfectness and virtual perfectness.** Throughout this section we have usually made the assumption that  $G = G_W$ . When  $W = \gamma_n$  for some  $n$ , this is equivalent to the statement that  $G$  is perfect. However it is evident that in most arguments it is sufficient to replace any given element  $g$  with a (fixed) power  $g^m$ . In particular, the theorems in this section remain true under the weaker hypothesis that for every  $g \in G$  there is a positive integer  $m$  such that  $g^m \in G_W$ ; equivalently, the quotient  $G/G_W$  is torsion. For general  $W$  this is implied by, but weaker than, the condition that  $G_W$  has finite index in  $G$ , even if  $G$  is finitely generated. But for  $W = \gamma_n$ , the two conditions are equivalent when  $G$  is finitely generated.

**Lemma 3.33.** *For any  $n$ , the quotient  $G/G_n$  is torsion if and only if  $G/G_2$  is torsion.*

*Proof.* One direction is obvious. We prove the other direction. Suppose  $G/G_2$  is torsion, and suppose by induction we have shown  $G/G_k$  is torsion for some  $k \geq 2$ . Let  $g \in G_k$ , and write  $g = [a_1, b_1][a_2, b_2] \cdots [a_r, b_r]$  where each  $b_r \in G_{k-1}$ . Let  $m$  be such that  $b_i^m \in G_k$  for all  $i$ . Then

$$\begin{aligned} g^m &= [a_1, b_1]^m \cdots [a_r, b_r]^m \mod G_{k+1} \\ &= [a_1, b_1^m] \cdots [a_r, b_r^m] \mod G_{k+1} \\ &= 0 \mod G_{k+1} \end{aligned}$$

By induction, the lemma is proved.  $\square$

Thus, Theorem 3.25 remains true with “perfect” replaced by “virtually perfect” (or even by: “group whose abelianization is torsion”).

Given a group  $G$ , one may attempt to obtain a lower bound on  $\text{sl}(*|\gamma_n)$  on elements of  $G_n$  by embedding  $G$  in a perfect group  $H$ , and using Theorem 3.23 applied to  $H$  together with monotonicity of  $\text{sl}(*|*)$  under homomorphisms. It is sometimes easier to take  $H$  to be virtually perfect rather than perfect; we shall see an example in § 4.

#### 4. GROPEs

The purpose of this section is to show how hyperbolic geometry can be used to give straightforward lower bounds on  $\text{l}(*|\gamma_3)$ . We give some examples that show that the uniform comparisons in Theorem 3.25 and Proposition 3.27 for *stable* lengths do not have any analogue for *unstable* length.

The arguments depend on the geometry of certain objects called *grope*s. We do not discuss here the most general kind of grope, but only the simplest nontrivial examples. For a general introduction to grope, see § 13 of [9]. Informally speaking, grope topologize the commutator calculus, and questions about expressing elements in groups as products of  $\gamma_n$  (or  $\beta_n$ ) words can be translated into questions about the existences of maps of certain kinds of grope to spaces. We also use some elementary facts from the theories of  $\text{CAT}(-1)$  complexes and pleated surfaces. A basic reference for the first is [2], especially pp. 347–362. A basic reference for the second in [22], Chapter 8.

Let  $S_n$  be an oriented surface of genus  $n$  with one boundary component. Let  $S_{n,1}$  be obtained by attaching a once-punctured torus to each of a maximal collection of pairwise disjoint homologically essential loops  $\beta_i$  in  $S_n$  (there are  $n$  such). We denote by  $\partial S_{n,1}$  the boundary of  $S_{n,1}$  (contained in  $S_{n,1}$ ).

Now, let  $G$  be a group, and let  $X$  be a space with  $\pi_1(X) = G$ . If  $g \in G$  is given, let  $\gamma : S^1 \rightarrow X$  be a loop whose free homotopy class corresponds to the conjugacy class of  $g$ . Observe that from the definitions,  $\text{l}(g|\gamma_3) \leq n$  if and only if there is a map  $f : S_{n,1} \rightarrow X$  such that  $\partial f : \partial S_{n,1} \rightarrow X$  factors through a homeomorphism  $h : \partial S_{n,1} \rightarrow S^1$  in such a way that  $\gamma \circ h = \partial f$ . Informally, the  $\gamma_3$ -length of  $g$  is at most  $n$  if and only if there is a map from  $S_{n,1}$  to  $X$  whose boundary wraps around  $\gamma$ .

**Proposition 4.1.** *For each  $n$ , let  $F$  be free on the generators  $x_1, y_1, \dots, x_n, y_n, z$  and let  $w_n = [z, [x_1, y_1][x_2, y_2] \cdots [x_n, y_n]]$ . Then  $\text{l}(w_n|\gamma_3) \geq 2n/3$ .*

*Proof.* We build a  $K(F, 1)$  (called  $X$ ) as follows. Start with a hyperbolic once-punctured torus  $S$  with totally geodesic boundary. Let  $\alpha$  be an embedded geodesic

in  $S$  representing the meridian. Take another hyperbolic surface  $S'$  of genus  $n$  with totally geodesic boundary of length  $\text{length}(\alpha)$ , and attach  $\partial S'$  isometrically to  $\alpha$ . The resulting space  $X$  is a  $K(F, 1)$ , and the conjugacy class of  $w_n$  is represented by the boundary circle  $\partial S$ , which by abuse of notation, we denote  $\partial X$ .

Since  $X$  is obtained by gluing convex hyperbolic 2-complexes along convex subsets, it is itself a  $\text{CAT}(-1)$  2-complex. If  $l(w_n|\gamma_3) = m$ , there is a map  $f : S_{m,1} \rightarrow X$  sending  $\partial S_{m,1}$  to  $\partial X$ . We homotop this map to a *pleated representative* in a special way. First we choose an ideal triangulation of  $S_m$  for which the geodesic representatives of the  $\beta_i$  (with notation as above) are contained in the pleating locus  $\lambda$ . The map  $f$  can be homotoped on  $S_m$  to take each leaf of  $\lambda$  to a geodesic in  $X$ , and to be 1-Lipschitz on each ideal triangle of  $S_m - \lambda$  for some hyperbolic metric on  $S_m$ . Then  $f$  can be homotoped rel.  $\cup_i \beta_i$  to a pleated representative with respect to some hyperbolic metric on each once-punctured torus component of  $S_{m,1} - S_m$ .

The key property of a pleated map is that it is area non-increasing. Moreover, it is surjective onto  $X$ , since for any point  $p \in X$ , the conjugacy class of  $w_n$  is not in  $[\pi_1(X - p), [\pi_1(X - p), \pi_1(X - p)]]$ . By Gauss–Bonnet,  $\text{area}(S_{m,1}) = 2\pi \cdot (3m - 1)$  and  $\text{area}(X) = 2\pi \cdot 2n$ . Hence  $3m - 1 \geq 2n$  and therefore  $m \geq 2n/3$ , as claimed.  $\square$

On the other hand,  $sl(w_n|\gamma_3) \leq 1$  for all  $n$ , by Proposition 3.27. This example shows that the comparison theorem (Theorem 3.25) has no analogue for unstable  $\gamma_n$ -lengths, even if  $n = 3$ .

## REFERENCES

- [1] C. Bavard, *Longueur stable des commutateurs*, Enseign. Math. (2) **37** (1991), no. 1-2, 109–150
- [2] M. Bridson and A. Haefliger, *Metric spaces of non-positive curvature*, Grundlehren der math. Wiss. **319**, Springer-Verlag, Berlin 1999
- [3] D. Burago, S. Ivanov and L. Polterovich, *Conjugation-invariant norms on groups of geometric origin, Groups of diffeomorphisms*, 221–250, Adv. Stud. Pure Math., 52, Math. Soc. Japan, Tokyo, 2008.
- [4] M. Burger and N. Monod, *Bounded cohomology of lattices in higher rank Lie groups*, J. Eur. Math. Soc. (JEMS) **1** (1999), no. 2, 199–235
- [5] D. Calegari, *Stable commutator length in subgroups of  $\text{PL}^+(I)$* , Pacific J. Math. **232** (2007), no. 2, 257–262
- [6] D. Calegari, *scl*, MSJ Memoirs **20**, Mathematical Society of Japan, Tokyo 2009
- [7] D. Calegari, *Quasimorphisms and laws*, Algebr. Geom. Topol. **10** (2010), no. 1, 215–217
- [8] D. Calegari and D. Zhuang, *Large scale geometry of commutator subgroups*, Algebr. Geom. Topol. **8** (2008), no. 4, 2131–2146
- [9] J. Cannon, *The recognition problem: what is a topological manifold?* Bull. Amer. Math. Soc. **84** (1978), no. 5, 832–866
- [10] M. Culler, *Using surfaces to solve equations in free groups*, Topology **20** (1981), no. 2, 133–145
- [11] A. Gleason, *Groups without small subgroups*, Ann. of Math. (2) **56** (1952), 193–212
- [12] M. Gromov, *Volume and bounded cohomology*, IHES Publ. Math. **56** (1982), 5–99
- [13] M. Gromov, *Asymptotic invariants of infinite groups*, Geometric group theory, Vol. 2 (Sussex, 1991), 1–295, London Mathematical Society LNS **182**, Cambridge University Press, Cambridge 1993
- [14] D. Kotschick, *Quasi-homomorphisms and stable lengths in mapping class groups*, Proc. Amer. Math. Soc. **132** (2004), no. 11, 3167–3175
- [15] M. Liebeck, E. O’Brian, A. Shalev and P. Tiep, *The Ore conjecture*, J. Eur. Math. Soc. (JEMS) **12** (2010), no. 4, 939–1008
- [16] D. Montgomery and L. Zippin, *Small subgroups of finite-dimensional groups*, Ann. of Math. (2) **56** (1952), 213–241
- [17] H. Neumann, *Varieties of groups*, Springer-Verlag, New York 1967



- [18] A. Rhemtulla, *A problem of bounded expressibility in free products*, Proc. Camb. Phil. Soc. **64** (1968), 573–584
- [19] D. Segal, *Words: notes on verbal width in groups*, London Mathematical Society LNS **361**, Cambridge University Press, Cambridge 2009
- [20] A. Shalev, *Word maps and Waring type problems*, J. Amer. Math. Soc. **22** (2009), no. 2, 437–466
- [21] A. Shalev, *Word maps, conjugacy classes, and a noncommutative Waring-type theorem*, Ann. of Math. (2) **170** (2009), no. 3, 1383–1416
- [22] W. Thurston, *Geometry and Topology of 3-manifolds*, Notes from lectures at Princeton University, available from the MSRI, circa 1979; a.k.a. “Thurston’s Notes”
- [23] D. Witte-Morris, *Bounded generation of  $SL(n, A)$  (after D. Carter, G. Keller, and E. Paige)*, New York J. Math. **13** (2007), 383–421

DEPARTMENT OF MATHEMATICS, CALTECH, PASADENA CA, 91125

*E-mail address:* dannyc@its.caltech.edu

DEPARTMENT OF MATHEMATICS, VANDERBILT UNIVERSITY, NASHVILLE TN, 37240

*E-mail address:* dongping.zhuang@vanderbilt.edu